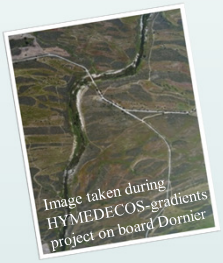
The EUFAR Data Archive

Wendy Garland - Centre for Environmental Data Archival (CEDA), STFC, UK - wendy.garland@stfc.ac.uk

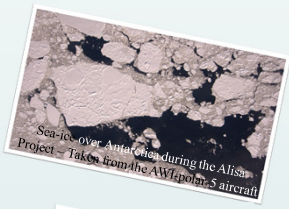


What is EUFAR?

The EUropean Facility for Airborne Research in Environmental and Geo-sciences (EUFAR2) (2014-2017) is an Integrating Activity of the 7th Framework Programme (FP7) of the European Commission, following three previous contracts under FP5, FP6 and FP7, bringing together 24 European institutions and organisations involved in airborne research, operating 18 instrumented research aircraft and providing access to a further 3 remote-sensing instruments.

Airborne and remote sensing measurements are an important source of data for developing the understanding of our atmosphere and its processes, and for validating meteorological and climate models and satellite instruments. Many European countries operate research aircraft but users are confined to the instruments and limitations (altitude, endurance) of their national facilities.

EUFAR's goal is to provide easy and open access to the airborne research facilities that are most suited to their needs by offering Transnational Access (TA) to national infrastructures, to reduce redundancy, fill the gaps, and optimise the use and development of the airborne infrastructure. It promotes the use of our research aircraft, especially for young scientists from countries where such facilities are lacking, by providing education and training courses on airborne research. EUFAR aims to improve the quality of the service by strengthening expertise through knowledge exchange, development of standards and protocols, provide access to the data collected, and engage joint instrumental research activities. It also supports innovation in airborne research, and promotes a culture of cooperation between EUFAR experts and SMEs to transfer airborne research instruments, methodologies and software into new products.



EUFAR Data Archive

A key component of the EUFAR project is the **central gateway to the data and supporting metadata** collected by the instrumented aircraft participating in EUFAR TA projects and training courses.

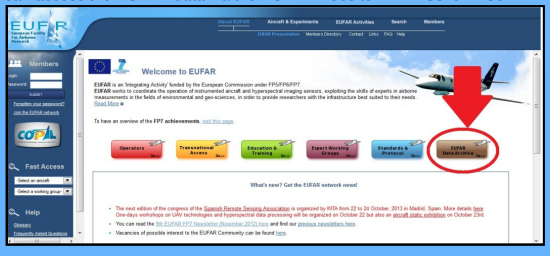
At the start of the EUFAR FP7 project (2008) only two EUFAR aircraft routinely made their data available through accessible online archives. Other data was passed straight to the commissioning scientists in formats convenient to themselves at the time of production. Discovering and acquiring this data was difficult outside of the original teams and therefore re-use was limited. The EUFAR data archive connects existing distributed archives and a provides a dedicated repository for data that would otherwise have been offline.

This portal makes all EUFAR data discoverable and widely accessible online through a single gateway. Data are stored in community-agreed data formats with appropriate well-documented metadata to enable re-use without recourse to the data provider. The inclusion of the archive within the British Atmospheric Data Centre (BADC) part of the Centre for Environmental Data Archival (CEDA) (www.ceda.ac.uk) enables it to be managed efficiently and consistently by experienced aircraft data scientists and to benefit from the latest curation best-practices. Dedicated data-provider and user support from experienced data archiving staff are also available through the BADC helpdesk. The EUFAR data is stored in CEDA's JASMIN petascale processing and storage super-data-cluster.

CEDA's JASMIN petascale processing and storage super-data-cluster
(photo Jonathan Churchill)

All EUFAR data is publicly available*. Web links were made to the existing aircraft archives (within the CEDA family) making all EUFAR data visible though a single gateway. *(a registration is required to track usage – but is open to all)

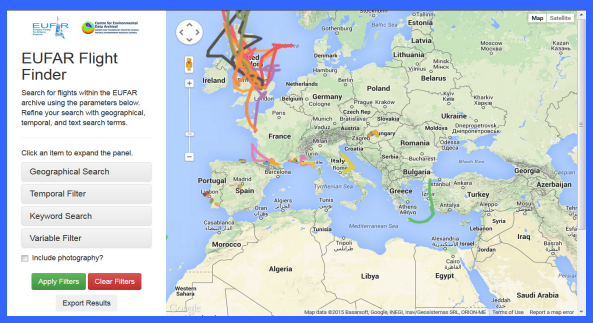
You can access the EUFAR data via the EUFAR website **www.eufar.net**



(or directly at badc.nerc.ac.uk/data/eufar/)

EUFAR Flight Finder Tool

A geospatial/ temporal/keyword search tool is being developed to enable detailed searches for flights within the EUFAR data archive. Location information, flight tracks, temporal information, along with parameter and key words have been extracted from the data and are searchable using elastic search. The results are displayed on a map interface with links to the data in the archive, and visualisation tools



Formats

Data are stored in community agreed standard data formats with supporting metadata to facilitate reuse. The use of standard formats facilitates wider access to the data and enables general tools to be used and developed.

Priority data formats are:

- CF-NetCDF* for in-situ measurements
- HDF and the standard binary format, ENVI-BIL for hyperspectral data.

(*CF =conventions for climate and forecast metadata http://cf-convention.github.io/)

Support to data providers

To achieve this fully populated central archive of formatted data requires dedicated support and advice by the BADC to each aircraft instrument team. This includes help with:

- Achieving the agreed standard data formats
- Identifying CF-compliant standard metadata names
- Filenaming conventions
- Data transfers – including large volumes of data

Achievements

- Dedicated online archive set up for otherwise-offline aircraft data.
- Connections to existing online aircraft archives ensuring all EUFAR data available to view and download from the same location.
- 9.2TB data from 154 flights/flightlines flights/flightlines made by 12 aircraft archived and accessible for posterity
- 42 separate flight campaigns supported plus community data sharing following Eyjafjallajökull volcanic eruption in 2010
- Dedicated individual formatting and upload support given to all data providers.
- All data publicly available via www.eufar.net or directly at badc.nerc.ac.uk/data/eufar/

Science & Technology Facilities Council

British Atmospheric Data Centre
NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL

Centre for Environmental Data Archival
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

SEVENTH FRAMEWORK PROGRAMME

www.eufar.net